Analysing movement data in a dynamic geographic context

Jean Damascène Mazimpaka

PhD student, Geoinformatics Group, University of Augsburg

Email: jean.mazimpaka@geo.uni-augsburg.de

Recent advances in location-aware technologies have provided an opportunity to learn the characteristics of moving entities and locations from traces of the moving entities. In the framework of analysing movement data, researchers have recently realised the need to analyse the traces in combination with their embedding context. In general, considering the context supports interpretation of the movement patterns discovered. In particular, linking the movement to a geographic context, which may affect it either by enabling or limiting it, can support understanding why entities move the way they do. The geographic context describes the movement situation in reference to the earth surface. Some work has been done to consider a static geographic context. For example, the embedding geographic space (e.g., the road network) and static objects (e.g., points of interest) have been considered. However, the work integrating a dynamic geographic context (i.e., a context changing with time) in movement data analysis is still limited and we lack a comprehensive approach to this integration.

The aim of my PhD thesis is to develop a comprehensive approach for integrating a dynamic geographic context into the analysis of movement data. The integration of a dynamic geographic context into the analysis of movement data presents the following challenges, which frame the research questions that I address. Firstly, a geographic context presents a lot of diversities which need to be identified along with their specific requirements. For example, while a football match in a stadium can be considered to be a zero dimensional object (represented as a point) with respect to a car moving on the road network of a city, a road extends over multiple locations. For the case of a road, there are spatial dynamics which need to be explicitly considered in addition to temporal dynamics. Secondly, the temporal dynamics and spatial dynamics need to be considered while linking the movement to a dynamic geographic context. Thirdly, the analysis process comprises several steps from pre-processing to pattern interpretation. Therefore, an appropriate step at which the context data should be integrated needs to be identified. In addition, some analysis operations have to be proposed.

The challenge of diversity has been addressed by establishing a classification of context elements based on their properties. I distinguished *static objects*, *dynamic objects*, *events*, and *space*. Static objects are objects with a fixed location (*e.g.*, a school located in the vicinity of a road on which a car is moving). Dynamic objects are other moving objects (*e.g.*, a pedestrian crossing the road segment in which a car is moving). Events are spatial happenings with a bounded time interval (*e.g.* a concert in a theatre located in the vicinity of a road on which a bus is moving*)*.The space is a set of geographic locations where the entity can be found during its movement (*e.g.*, a road on which a car is moving).

In order to link the context to movement I adopted an approach based on the change of spatial relations between the moving entity and the context element over time. This change has been called *movement interaction.* Movement interactions were formally defined based on the change of distance relations. I proposed a small set of basic interactions (*approaching, arriving, stopping, leaving,* and *moving-away*) that apply on simple cases of context elements that are static and zero-dimensional. This interaction set was then extended to context made of moving objects, and finally to context made of context elements of type "space". The interactions were organised in conceptual neighbourhood graphs, which can further support a qualitative reasoning on relations between moving entities and the dynamic geographic context.

The analysis of movement data and related context data starts by extracting these movement interactions. Data mining and spatial analysis and methods are used to extract interactions because interactions are movement patterns but which embed context information. The interactions have attributes such as the start and end times, and the identifiers of the moving entity and the context element involved. The next step after the extraction of interactions is the analysis of extracted interactions. To this end, the interactions are quantified and the variation of dynamic attributes associated with the movement explored for detecting any possible correlation with or dependency on the dynamics of the context. I proposed different analysis operations based on statistical methods (*e.g.*, aggregation, and correlation coefficient computation) and data mining methods (*e.g.,* sequence analysis). Aggregation methods allow handling both spatial and temporal dynamics of the geographic context. Sequence analysis exploits the link that exists between interactions as expressed by the conceptual neighbourhood graphs.

Considering that the analysis framework proposed follows the Knowledge Discovery in Database (KDD) process, I identified two potential steps of the analysis process where context data can be integrated. These are the data mining (or pattern discovery) step and the pattern interpretation step. Context data have been integrated into the analysis at these two steps in different experiments in order to identify the appropriate step. The experiments showed that the integration of context data at the pattern discovery step is the most appropriate because it avoids the discovery of patterns that cannot be interpreted due to missing related context data. Moreover, some movement patterns (*e.g.*, group patterns) can be discovered only if required context data, in the form of trajectories of other moving entities, are available.

The proposed analysis framework is being evaluated on real data of bus and taxi movement in urban area and context data such as road network and big social events. The evaluation is done by executing a number of experiments each focussing on a specific aspect of the challenges being addressed. The analysis framework is general for use in different application domains. In particular, the experiments executed show that the combined consideration of movement data and context elements can be usefully applied in urban planning and transportation. For example, linking the human mobility to the dynamics of traffic congestion on road segments can improve the prediction and evolution of traffic jams and enhance journey planning applications. Likewise, understanding the effect of big events on the movement of public transportation vehicles as well as the temporal pattern of arrival and departure of event attendees can support the management of transportation in cases of big events.